

How to Axiomatize School Geometry

Eliahu Levy

Department of Mathematics

Technion – Israel Institute of Technology, Haifa 32000, Israel

email: eliahu@techunix.technion.ac.il

This is an attempt to present axioms for Euclidean Geometry, aiming at the following goals:

- To work with “geometrical” notions. Thus we would not merely identify points in the plane with pairs of real numbers, which means that a particular coordinate system is given special status.
- To be appropriate to the way geometry is done in higher mathematics (including physics and engineering). This means that the algebraic nature of much of geometry need not be hidden.
- To respond to the desire that one would confidently accept empirically/intuitively that the axioms are valid in our physical everyday world (or rather in the usual idealization that “geometry” is). This seems to disfavor taking the Theorem of Pythagoras as an axiom.
- To have accessible the usual rigor of “pure” mathematics, and to make the axioms satisfying by the standards of the latter. In particular, not to take as an axiom something that can be naturally proved. Note that some “topological” notions, necessary for the rigor of the presented axioms, can be readily kept silent with an “unsophisticated” audience (such as school).

The style in the sequel is intended for those accustomed to mathematical writings, in order to make the mathematical contents clear. Of course, in case an approach in this spirit can be practiced in school the style of presentation must be quite different.

1 The Axioms: Plane Geometry

Primitive notions:

- A set P (The **Plane**), whose elements are called **points**. For the sophisticated – the Plane is assumed a Hausdorff topological space. We shall let x, y, z, a, b, c etc. vary over points.

- A relation among 3 points, indeed a commutative “algebraic” operation: z is the middle between x and y (to be written $z = \text{Middle}(x, y)$).
- An equivalence relation on $P \times P$: two pairs of points have the same distance.

We shall have altogether three axioms.

By \mathbf{R} we denote, as usual, the set of **real numbers** (for the unsophisticated - just the set of **numbers**, representable, say, as possibly unending decimal fractions).

Axiom 1 (Axiom of Coordinates). *There is a bijection (coordinates system) between \mathbf{R}^2 and the Plane (which is a homeomorphism) such that any mapping given in the coordinates by $x \mapsto a + x$ or $x \mapsto a - x$ (here $a \in \mathbf{R}^2, x \in \mathbf{R}^2$) “preserves the geometry”: it is an “isomorphism” with respect to the middle operation and maps any pair of points to a pair with the same distance.*

Note that we did not take the coordinates as a primitive notion: there may be many such bijections, the axiom saying that there is at least one.

The empirical/intuitive evidence for this axiom is plain: one encounters such coordinates daily (with the “rough” everyday correlate of the idealized “set of points”). Maybe it is more intuitive to postulate the stronger requirements that the reflections $(x_1, x_2) \mapsto (\alpha \pm x_1, x_2)$ and $(x_1, x_2) \mapsto (x_1, \alpha \pm x_2)$ ($\alpha \in \mathbf{R}$) preserve the geometry, the existence of enough reflections being implicit even in Euclid.

As a simple consequence we may prove

Theorem 1 *In any coordinates system satisfying Axiom 1 the middle operation $\text{Middle}(a, b)$ corresponds to the “algebraic middle” $\frac{a+b}{2}$ $a, b \in \mathbf{R}^2$.*

Proof The map $x \mapsto (a+b) - x$ maps $a \mapsto b, b \mapsto a$, and preserves the geometry. Hence it fixes $\text{Middle}(a, b)$. But its only fixed point is $\frac{a+b}{2}$.

QED

Define, for integer $n \geq 2$, an **n -ruler** as a sequence of points $(a_i)_{0 \leq i \leq n}$ such that for any $0 < i < n$ $a_i = \text{Middle}(a_{i-1}, a_{i+1})$

Theorem 1 implies that for any coordinates system as in Axiom 1, a ruler is just an “algebraic ruler”, i.e. a sequence with constant difference. We may deduce:

Fact *For any integers $n \geq k > l \geq 0$ and points a, b there is a unique n -ruler $(c_i)_{0 \leq i \leq n}$ with $a = c_k$ and $b = c_l$.*

If we define, with respect to some coordinates system, an “algebraic **straight line**” as usual (as a set $L \subset \mathbf{R}^2$ of the form $L = \{a + \lambda c \mid \lambda \in \mathbf{R}\}$ where $a, c \in \mathbf{R}^2, c \neq 0$) then the straight line joining $a, b \in P$ contains the “rational line” joining a and b , i.e. the set of all

points obtained by constructing n -rulers according to the Fact, and is its closure. Thus the notion of straight line is independent of the coordinates (note that we needed the topology here, and that just 2-rulers would have sufficed).

A quadrangle (a, b, c, d) is an “algebraic parallelogram” with respect to some coordinates system if $a - b = d - c$. But this is equivalent to (a, c) and (b, d) having the same “algebraic middle”, i.e. to $\text{Middle}(a, c) = \text{Middle}(b, d)$. Thus the notion of parallelogram is again “geometrical” – independent of the coordinates system. This allows us to define the **vectors** geometrically as “differences of pairs of points”, that is, say, as equivalence classes of pairs of points by the equivalence relation defined by parallelograms. (Thus a point minus a point is a vector, and a point plus a vector is a point). Any coordinates system lets us identify the vectors with \mathbf{R}^2 , thus making them into a 2-dimensional \mathbf{R} -vector space, and one easily shows that the vector operations can be defined “geometrically” – independent of the coordinates. (For multiplication by general real numbers we again need the topology). Denote the 2-dimensional space of vectors by V . By $\text{End}(V)$ we will mean the space of linear self-maps of V .

In so far we had little to do with the primitive equivalence relation of two pairs of points having the same distance. Now we come to it. By the requirements from coordinates in Axiom 1 any two pairs with the same vector difference have the same distance, thus we get an equivalence relation between vectors: **having the same length**, and moreover v and $-v$ always have the same length.

Define an **isometry** as an invertible linear self-map $\mathcal{U} \in \text{End}(V)$ mapping each vector into a vector with same length. The set of isometries is a group. By the above, -1 belongs to this group. (Here and in the sequel we identify a scalar operator with the scalar).

The two remaining axioms deal with isometries. They have a markedly algebraic flavor, which seems justifiable in view of the above.

Axiom 2 (Axiom of Isotropy). *The group of isometries is transitive on a set of all vectors of the same length, and is also transitive on the set of 1-dimensional subspaces of V . (That is: for any two vectors of the same length, or two 1-dimensional subspaces \exists an isometry mapping one to the other).*

Instead of the first half of Axiom 2, one could take the group of isometries as a primitive notion and define vectors to have the same length iff an isometry maps one to the other.

Axiom 3 (Axiom of Boundedness). *The group of isometries is bounded (as a subset of the 4-dimensional \mathbf{R} -vector space $\text{End}(V)$).*

Axiom 2 is related to the empirical/intuitive possibility of motions (rotations etc.), which is often expressed by congruence axioms. Axiom 3 postulates that the circle is bounded in a coordinate system, in spite of the latter extending to infinity in the idealization which is “geometry”.

Now we shall be able to use the following theorem from algebra/analysis to obtain that there is a positive-definite quadratic form Q on V such that vectors $u, v \in V$ have the same

length iff $Q(u) = Q(v)$ (thus we have the Theorem of Pythagoras). The resort to such theorem here seems natural from our point of view. Unfortunately, proving it requires some mathematical sophistication.

Theorem 2 *For any bounded group $G \subset GL(V)$, where V is a 2-dimensional \mathbf{R} -vector space, there exists a G -invariant positive-definite quadratic form Q .*

We give three proofs, differing in the tools used.

Proof 1 \overline{G} is a compact group, thus admits a normalized Haar measure μ . Take any positive-definite quadratic form Q_0 and take as Q the average

$$Q(v) = \int_{g \in \overline{G}} Q_0(gv) d\mu(g).$$

QED

This proof works for any finite-dimensional V over \mathbf{R} .

Proof 2 This again works for any finite-dimensional V .

Let W be the \mathbf{R} -vector space of quadratic forms on V , and W_+ the set of the positive-definite ones (this set is an open convex cone). G acts on W in the canonical way: $(gQ)(v) := Q(g^{-1}v)$, $Q \in W$, and leaves W_+ invariant.

Choose a norm $|||_0$ on W , say the maximum of the absolute values of the matrix entries with respect to a basis of V . Replace $|||_0$ by the G -invariant norm

$$\|Q\| := \sup_{g \in G} \|gQ\|_0.$$

We know that there is a fixed integer $N > 0$ such that any subset of W with $|||$ -diameter $\leq d$ can be covered by at most N sets of $|||$ -diameter $\leq \frac{d}{2}$. If $K \subset W$ is bounded non-empty G -invariant convex, say the convex hull of the orbit of some Q , and $\text{diam}(K) \leq d$, then we have a finite set $F \subset K$, of at most N elements, such that $\forall Q \in K \exists Q' \in F \|Q - Q'\| \leq \frac{d}{2}$.

This holds, in particular, for any Q of the form gQ_0 , $g \in G$, $Q_0 := \frac{\sum F}{\#F}$. Thus gQ_0 has distance $\leq \frac{d}{2}$ from some $Q' \in F$ and distance $\leq d$ from the other members of F . This implies $\|gQ_0 - Q_0\| < \gamma d$ where $\gamma := \frac{2N-1}{2N} < 1$. Since $|||$ is G -invariant, we have that the orbit of Q_0 , hence its convex hull, has diameter $\leq \gamma d$.

So we know that any bounded non-empty G -invariant convex set $\subset W$ with diameter $\leq d$ has a non-empty G -invariant convex subset of diameter $\leq \gamma d$. Repeating the process we get an infinite sequence of nested sets which converges to a G -invariant $Q_I \in W$. If we ensure that for any $Q \in K$ $Q(v) \geq \alpha \|v\|^2$ for some fixed norm $|||$ on V and some fixed $\alpha > 0$, then Q_I will be positive-definite.

QED

Proof 3 This is a purely algebraic proof, using $\dim(V) = 2$.

There is only a 1-dimensional space of antisymmetric forms on V ; that is, a choice of such non-zero form, which we make and denote by $u \wedge v$ $u, v \in V$, is possible and unique up to a scalar multiple. (This follows from $b_1 \wedge b_1 = 0$, $b_2 \wedge b_2 = 0$, $b_2 \wedge b_1 = -b_1 \wedge b_2$ which any such form must satisfy for a basis (b_1, b_2) , while these formulas indeed give a non-zero antisymmetric form.)

The determinant and trace of the matrix of an $\mathcal{A} \in \text{End}(V)$ are independent of the basis (since different bases give similar matrices), hence we may speak of $\det(\mathcal{A})$ and $\text{tr}(\mathcal{A})$. The characteristic polynomial of \mathcal{A} is

$$(1) \quad x^2 - \text{tr}(\mathcal{A}) \cdot x + \det(\mathcal{A}),$$

its real roots are the real eigenvalues of \mathcal{A} , and plugging \mathcal{A} in it gives 0, by the Cayley-Hamilton Theorem.

For a traceless $\mathcal{A} \in \text{End}(V)$ (i.e. with $\text{tr}(\mathcal{A}) = 0$), we obtain from the Cayley-Hamilton Theorem that \mathcal{A}^2 is a scalar, equal to $-\det(\mathcal{A})$, and of course to $\frac{1}{2}\text{tr}(\mathcal{A}^2)$. This scalar gives a quadratic form on the 3-dimensional \mathbf{R} -vector space $\{\mathcal{A} \in \text{End}(V) \mid \text{tr}(\mathcal{A}) = 0\}$, where the corresponding symmetric bilinear form is

$$\langle \mathcal{A}, \mathcal{B} \rangle := \frac{1}{2}\text{tr}(\mathcal{A}\mathcal{B}) = \frac{1}{2}\text{tr}(\mathcal{B}\mathcal{A}).$$

Checking the orthogonal basis $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ shows that the signature is $(+, +, -)$. Thus there cannot be two orthogonal elements with non-positive value of the quadratic form.

If $G \subset \text{GL}(V)$ is a bounded group, then the image of G by \det is a bounded subgroup of $\mathbf{R}^\times (= \mathbf{R} \setminus \{0\})$, therefore $\det(\mathcal{A}) = \pm 1$ for $\mathcal{A} \in G$. Also, if $\mathcal{A} \in G$ has a real eigenvalue λ , then $\mathcal{A}^n \in G$ for integer n and has the eigenvalue λ^n , and these must be bounded, therefore $\lambda = \pm 1$.

Hence if $\mathcal{A} \in G$ has determinant 1, (1) cannot have a real root different from ± 1 , which implies $|\text{tr}(\mathcal{A})| \leq 2$. For such \mathcal{A} , (1) can be written as (recall that $\det(\mathcal{A}) = 1$):

$$(2) \quad \left(x - \frac{1}{2}\text{tr}(\mathcal{A})\right)^2 = -\left(1 - \left(\frac{1}{2}\text{tr}(\mathcal{A})\right)^2\right) \leq 0.$$

By Cayley-Hamilton, such \mathcal{A} can be written as $\frac{1}{2}\text{tr}(\mathcal{A}) + \mathcal{J}$ where \mathcal{J} is traceless with $\mathcal{J}^2 = -\det(\mathcal{J})$ non-positive, equal to $-\left(1 - \left(\frac{1}{2}\text{tr}(\mathcal{A})\right)^2\right)$. Also, the product of $\frac{1}{2}\text{tr}(\mathcal{A}) \pm \mathcal{J}$ is 1, thus $\mathcal{A}^{-1} = \frac{1}{2}\text{tr}(\mathcal{A}) - \mathcal{J}$.

If we had $\mathcal{J}^2 = 0$ without $\mathcal{J} = 0$, then $\frac{1}{2}\text{tr}(\mathcal{A}) = \pm 1$, thus $\pm \mathcal{A} = 1 + \mathcal{J}_1$, $\mathcal{J}_1^2 = 0$, $\mathcal{J}_1 \neq 0$ and $(\pm \mathcal{A})^n = 1 + n\mathcal{J}_1$ contradicting the boundedness of G . Hence if \mathcal{A} is not ± 1 then $\mathcal{J}^2 < 0$ and $|\text{tr}(\mathcal{A})| < 2$.

We claim that if $\mathcal{A}_1 = \tau + \mathcal{J}_1$ and $\mathcal{A}_2 = \tau + \mathcal{J}_2$, τ scalar, are elements of G with determinant 1 and the same trace, \mathcal{J}_1 and \mathcal{J}_2 being traceless with the non-positive square $-(1 - \tau^2)$, then $\mathcal{J}_2 = \pm \mathcal{J}_1$, that is $\mathcal{A}_2 = \mathcal{A}_1^{\pm 1}$. Indeed, in the above quadratic form on the space of traceless

elements of $\text{End}(V)$, given by the scalar square, \mathcal{J}_1 and \mathcal{J}_2 have the same non-positive square and thus are the sum and difference of the orthogonal $\frac{1}{2}(\mathcal{J}_1 \pm \mathcal{J}_2)$. These cannot both have negative square, the signature being $(+, +, -)$, and none can have square 0 if $\mathcal{J}_1 \neq \pm \mathcal{J}_2$. Hence in the latter case they have squares of strictly different signs which implies

$$|\langle \mathcal{J}_1, \mathcal{J}_2 \rangle| = \left| \frac{1}{2} \text{tr}(\mathcal{J}_1 \mathcal{J}_2) \right| > |\langle \mathcal{J}_i, \mathcal{J}_i \rangle| = 1 - \tau^2.$$

Returning to \mathcal{A}_1 and \mathcal{A}_2 this gives either $\mathcal{A}_1 \mathcal{A}_2$ or $\mathcal{A}_1 \mathcal{A}_2^{-1}$ is an element of G which has half-trace greater than 1 which we saw above is impossible.

Suppose that we have picked an $\mathcal{A} \in G$ with determinant 1 and \mathcal{A} not the scalar ± 1 . We have $\mathcal{A} = \frac{1}{2} \text{tr}(\mathcal{A}) + \mathcal{J}$, $\mathcal{A}^{-1} = \frac{1}{2} \text{tr}(\mathcal{A}) - \mathcal{J}$, \mathcal{J} is traceless and $\mathcal{J}^2 < 0$. Consider the symmetric bilinear form on V

$$(3) \quad B(u, v) := \frac{1}{2} [(\mathcal{A}u) \wedge v + (\mathcal{A}v) \wedge u] = (\mathcal{J}u) \wedge v.$$

If b_1 is a non-zero vector and $b_2 = \mathcal{J}b_1$, then b_2 cannot be $= \lambda b_1$, $\lambda \in \mathbf{R}$ because that would imply $\mathcal{J}^2 b_1 = \mathcal{J}b_2 = \lambda \mathcal{J}b_1 = \lambda b_2 = \lambda^2 b_1$. Therefore (b_1, b_2) is a basis, and $B(b_1, b_1) = b_2 \wedge b_1 \neq 0$. We have $B(b_1, b_2) = 0$ and $B(b_2, b_2) = (\mathcal{J}^2 b_1) \wedge (\mathcal{J}b_1) = -(\mathcal{J}^2)B(b_1, b_1)$. So we conclude that B or $-B$ must be positive-definite.

We claim that B is invariant under G . Indeed, for $\mathcal{B} \in G$:

$$\begin{aligned} B(\mathcal{B}u, \mathcal{B}v) &= \frac{1}{2} [(\mathcal{A}\mathcal{B}u) \wedge (\mathcal{B}v) + (\mathcal{A}\mathcal{B}v) \wedge (\mathcal{B}u)] = \\ &= \det(\mathcal{B}) \frac{1}{2} [(\mathcal{B}^{-1}\mathcal{A}\mathcal{B}u) \wedge v + (\mathcal{B}^{-1}\mathcal{A}\mathcal{B}v) \wedge u] \end{aligned}$$

But $\mathcal{B}^{-1}\mathcal{A}\mathcal{B}$ is an element of G with determinant 1 and the same trace as \mathcal{A} . By the above, it is equal to either \mathcal{A} or \mathcal{A}^{-1} and we find that $B(\mathcal{B}u, \mathcal{B}v)$ is one of $\pm B$. Since both are positive-definite or negative-definite, they are equal.

The theorem is hence proved except when all members of G with determinant 1 are scalars. If that is the case, then if there are no elements in G with determinant -1 we are done. In any case, for any $\mathcal{J} \in G$ with $\mathcal{J}^2 = -1$ we can, as above, construct a basis (b_1, b_2) , $b_2 = \mathcal{J}b_1$ and the matrix of \mathcal{J} in this basis is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ with determinant 1. Hence such a \mathcal{J} is excluded in our case, and all $\mathcal{A} \in G$ with $\det(\mathcal{A}) = -1$ must satisfy $\mathcal{A}^2 = 1$. Then for any $v \in V$ $v = \frac{1}{2}(v + \mathcal{A}v) + \frac{1}{2}(v - \mathcal{A}v)$ is a sum of eigenvectors of \mathcal{A} with eigenvalues 1 and -1 , respectively. Since $\det(\mathcal{A}) = -1$ we have a 1-dimensional space of each, i.e. \mathcal{A} has matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ in some basis. Any other member of G with determinant -1 is a multiple of this \mathcal{A} by a member of G with determinant 1, hence the only possible members of G are $\pm 1, \pm \mathcal{A}$ and one easily finds many positive-definite quadratic forms invariant with respect to these.

QED

One may remark, following Bourbaki, that after one has the quadratic form on V that determines equality of length, and thus a corresponding symmetric bilinear form $\langle \rangle$, one may

define a $\mathcal{J} \in \text{End } V$ by $\langle u, v \rangle = (\mathcal{J}u) \wedge v$ (see the third proof above) and prove that \mathcal{J}^2 is a negative scalar, hence by normalizing the \wedge one may have $\mathcal{J}^2 = -1$ (which determines \mathcal{J} up to sign). This turns V in a canonical way into a 1-dimensional complex vector space (didactically, the complex numbers may be defined by our “geometric” \mathcal{J} – in this approach every plane (say, in 3-space) has, strictly speaking, its own “complex numbers”.) Using this complex structure to do plane geometry is very fruitful. For example, angles (with their trigonometry) can be easily treated in a rigorous way.

2 Space Geometry

To axiomatize the Euclidean geometry of n -space, one may start from a set **Space** of points with exactly analogous primitive notions, replace \mathbf{R}^2 by \mathbf{R}^n in the Axiom of Coordinates, and postulate the Axioms of Isotropy and of Boundedness for every sub-2-plane of Space (or alternatively for Space itself). In passing from the existence of an equality-of-length – determining quadratic form on every plane to the existence of one such form for Space, one may use the well-known

Theorem 3 *Let V be an (not necessarily finite-dimensional) \mathbf{R} -vector space, $\dim(V) \geq 2$, and let $U \subset V$. If for every 2-dimensional subspace (=plane) $V' \subset V$ there exists a positive-definite quadratic form Q' on V' such that $U \cap V' = \{v \in V' \mid Q'(v) = 1\}$ then there exists a positive-definite quadratic form Q on the whole V such that $U = \{v \in V \mid Q(v) = 1\}$.*

Proof It is clear that the Q' are unique for each plane, and that they agree on intersections. Hence they define a function Q on V , positive on $V \setminus 0$ and it remains to prove that Q is quadratic, i.e. comes from a symmetric bilinear form on V . That form must be:

$$\langle u, v \rangle = \frac{1}{2}(Q(u+v) - Q(u) - Q(v)) \quad u, v \in V$$

and we have to prove that this is bilinear. Since we have $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$, Q being quadratic on the plane containing u and v , it remains to prove biadditivity. As we clearly have $\langle u, 0 \rangle = 0$, biadditivity will follow if we prove that $\langle u, v \rangle + \langle u, w \rangle$ depends only on u and $v+w$. This obtains from the following calculation (where one uses the parallelogram equality $2Q(a) + 2Q(b) = Q(a+b) + Q(a-b)$, $a, b \in V$, which holds since Q is quadratic in the plane containing a and b):

$$\begin{aligned} 2(\langle u, v \rangle + \langle u, w \rangle) &= Q(u+v) - Q(u) - Q(v) + Q(u+w) - Q(u) - Q(w) = \\ &= \frac{1}{2}(Q(2u+v+w) + Q(v-w)) - 2Q(u) - \frac{1}{2}(Q(v+w) + Q(v-w)) = \\ &= \frac{1}{2}(Q(2u+v+w) - Q(v+w)) - 2Q(u) \end{aligned}$$

QED